VIROLOGY 216, 165–173 (1996)
ARTICLE NO. 0043

Genetic Variants of Human T-Lymphotropic Virus Type II in American Indian Groups

ROBERT J. BIGGAR,*[,1] MARIA E. TAYLOR,† JAMES V. NEEL,‡ BRIAN HJELLE,§ PAUL H. LEVINE,*
FRANCIS L. BLACK,¶ GEORGE M. SHAW,† PAUL M. SHARP,‖ and BEATRICE H. HAHN†

*Viral Epidemiology Branch, National Cancer Institute, Bethesda, Maryland; †Departments of Medicine and Microbiology, University of Alabama at Birmingham, Birmingham, Alabama; ‡Department of Human Genetics, University of Michigan, Ann Arbor, Michigan; §Department of Pathology, University of New Mexico, Albuquerque, New Mexico; ¶Department of Epidemiology, Yale University Medical School, New Haven, Connecticut; and ‖Department of Genetics, University of Nottingham, Queens Medical Centre, Nottingham, United Kingdom

Received October 16, 1995; accepted December 8, 1995

The human T-lymphotropic virus type II (HTLV-II) is found in many New World Indian groups in North and South America and may have entered the New World from Asia with the earliest migration of ancestral Amerindians over 15,000 years ago. To characterize the phylogenetic relationships of HTLV-II strains infecting geographically diverse Indian populations, we used polymerase chain reaction to amplify HTLV-II sequences from lymphocytes of seropositive Amerindians from Brazil (Kraho, Kayapo, and Kaxuyana), Panama (Guaymi), and the United States (the Navajo and Pueblo tribes of the southwestern states and the Seminoles of Florida). Sequence analysis of a 780-base pair fragment (located between the *env* gene and the second exons of *tax*/*rex*) revealed that Amerindian viruses clustered in the same two genetic subtypes (IIa and IIb) previously identified for viruses from intravenous drug users. Most infected North and Central American Indians had subtype IIb, while HTLV-II infected members of three remote Amazonian tribes clustered as a distinct group within subtype IIa. These findings suggest that the ancestral Amerindians migrating to the New World brought at least two genetic subtypes, IIa and IIb. Because HTLV-II strains from Amazonian Indians form a distinct group within subtype HTLV-IIa, these Brazilian tribes are unlikely to be the source of IIa viruses in North American drug users. Finally, the near identity of viral sequences from geographically diverse populations indicate that HTLV-II is a very ancient virus of man.

INTRODUCTION

The human T-lymphotropic viruses (HTLV) types I and II are retroviruses with a worldwide distribution. HTLV-I appears to be endemic in many Old World indigenous peoples as well as immigrants to the New World and persons with contact to them (Poiesz *et al.,* 1993). HTLV-II has been most commonly found in North American and European intravenous drug users (IVDUs) (Hall *et al.,* 1994a). However, several recent reports have documented HTLV-II among New World Indians (summarized in Neel *et al.,* 1994). Tribes found to have endemic HTLV-II include the Seminoles of Florida, the Navajo and Pueblo tribes of the southwestern United States, the Maya of the Yucatan Peninsula, the Guaymi of Panama, the Wayu of Columbia, the Puma of Venezuela, the Ge-speaking peoples (Kraho and Kayapo) of central Brazil, and the Toba of Argentina. Ishak *et al.* (1995) recently described additional Brazilian tribes endemically infected with HTLV-II, and no doubt other Amerindian groups will be found to have endemic HTLV-II infections.

The observation that many tribes known to harbor HTLV-II have had no direct contact with each other for thousands of years led us to propose that HTLV-II was probably brought to the New World when ancestral Amerindians ("Paleoindians") first crossed the Bering Strait at least 15,000 years ago (Neel *et al.,* 1994). Endemic HTLV-II infection apparently also occurs in isolated groups in Africa (Goubau *et al.,* 1992; Gessain *et al.,* 1995), as well as in Mongolia (Hall *et al.,* 1994b), the latter representing a possible source of the HTLV-II found in today's Amerindian populations (Neel *et al.,* 1994).

In our previous studies of HTLV-II in Amerindians (Maloney *et al.,* 1992; Black *et al.,* 1994), samples were collected mainly in the 1960s and early 1970s, at a time when these Indians had limited contact with the outside world. Thus, it is likely that these subjects had few (if any) interactions, such as sexual relationships or receiving intravenous blood, that could have introduced HTLV into their population. Noteworthy, over 50% of older persons in some isolated groups were found to be infected, whereas none of the members of other tribes had HTLV-II infection (Maloney *et al.,* 1992; Black *et al.,* 1994). We interpreted this variation to be the result of founder effects, in which a small number of individuals who established new tribal groups in the past may not have had HTLV infection (Maloney *et al.,* 1992).

Recent phylogenetic studies of HTLV-II sequences report that there are two distinct subtypes, designated IIa

[1] To whom correspondence and reprint requests should be addressed at 6130 Executive Boulevard, EPN 434, Rockville, Maryland 20852. Fax: (301) 402-0817.

0042-6822/96 $12.00

and IIb, which infect IVDUs and Amerindians (Hall *et al.,* 1993; Switzer *et al.,* 1995a). With the exception of viruses from a few Navajo/Pueblo subjects (Switzer *et al.,* 1995a), most Amerindian HTLV-II strains thus far have been found to cluster within subtype IIb. Here we report sequence data showing that three different Brazilian Amerindian tribes are infected with HTLV-II belonging to a distinct subgroup within IIa. We also describe the phylogenetic relationships of HTLV-II strains infecting several widely separated Amerindian groups in Brazil, Panama, Florida, and the southwestern United States and discuss paradoxes relating to their epidemiology and molecular evolution.

MATERIALS AND METHODS

Amerindian populations

Blood samples were collected during field studies on Amerindian groups between 1966 and the present. Details are given in Table 1. Samples from the Seminoles and southwestern Indians have been widely circulated and have been tested in several laboratories. For example, SC, AG, and DSA have been studied here as well as by Ishak *et al.* (1995), although analyses were done on different genomic regions. For their identification, we include subjects' initials.

The recent histories of the groups studied may be instructive about tribal interactions. In South America, the original sample collections were undertaken between 1966 and 1987 to describe genetic features of Amerindian groups (Maloney *et al.,* 1992; Black *et al.,* 1994). Tribes were selected on the basis of their remoteness. The Kraho and Kayapo tribes of Tocantins and southern Para states in central Brazil speak two different Ge languages. One Brazilian subject was a Carib-speaking Kaxuyana. This tiny tribe illustrates the flow of tribal relationships. Thirty years ago, the tribe consisted of only 43 persons. In 1970, they relocated from just north of the Amazon to near the Surinam border, intermarried, and subsequently expanded in number. While different tribes must have had a common origin long ago, the distance between them is now extensive: the Kayapo are over 1000 kilometers from the Kraho, and the Kaxuyana are now over 700 kilometers from the Kayapo and 1,400 kilometers from the Kraho.

The Guaymi of Panama were sampled in 1974 from isolated villages on the north coast of Panama (Maloney *et al.,* 1992), which differs from other recently studied Guaymi, who had resettled from traditional villages to banana plantations within the last generation (Lairmore *et al.,* 1990) and were thus less isolated. The Navajo of New Mexico are linguistically Athapaskan (Na-dene) Indians and are thought to constitute one of the southernmost descendants of a later wave of migration (about 8000 years ago) across the Bering Strait. They have intermarried with Paleoindian tribes in the same area, loosely called the Pueblo tribes because of their dwellings (further attributions of tribal affiliation are not available). Both Navajo and Pueblo blood samples were collected in 1990 (Hjelle *et al.,* 1993). The Seminoles are the remnants of southeastern Amerindians who settled in the Everglades of Florida during the early 1800s to avoid persecution from early settlers. They have extensively intermarried with African-Americans living in the Everglades. Samples were collected in 1989–91 (Levine *et al.,* 1993).

Polymerase chain reaction

High molecular weight DNA was extracted as described (Maloney *et al.,* 1992) from uncultured peripheral blood mononuclear cells (PBMC) of seropositive Seminole, Guaymi, Pueblo, and Navajo subjects, uncultured leukocyte/erythrocyte pellets of seropositive Brazilian Amerindians, and cultured HTLV-II isolates derived from IVDUs originating from New Orleans, Miami, and Newark. Primer pairs were designed according to the published sequence of HTLV-II/MO (Shimotohno *et al.,* 1985) (see Table 2). PCR amplifications were carried out in 100 μl, containing 0.5–1 μg of genomic DNA, 10 m*M* Tris–HCl (pH 8.3), 50 m*M* KCl, 20 pmol of each primer, 20 μ*M* of each of the four deoxynucleoside triphosphates (dNTPs), 1.5 m*M* $MgCl_2$, 0.01% (wt/vol) gelatin, and 2.5 U of *Taq* polymerase. Samples were subjected to 30 amplification cycles.

As outlined in Fig. 1, different amplification strategies were pursued. Uncultured patient PBMCs were subjected to nested PCR analysis using primer pairs designed to amplify the region between *env* gene and the second exons of *tax*/*rex* (780 bp) either in one or two overlapping fragments (Fig. 1). Cultured HTLV-II was amplified in a single round of PCR. Amplification products were visualized by agarose gel electrophoresis, isolated from a preparative gel, purified with GeneClean (Bio 101 Inc., La Jolla, CA), and subcloned into M13 or pCR cloning vectors (Invitrogen, San Diego, CA) by T/A overhang. For control, we also amplified HTLV-II/MO from the C3-44 cell line.

DNA sequence analysis

Recombinant clones containing HTLV-II fragments were sequenced manually by the dideoxynucleotide chain termination method (Sequenase Kit; U.S. Biochemicals, Cleveland, OH). Generally, one clone per amplification product was sequenced, but from samples FF0754, AP3207, AP3208, AP3210, FF2085, F0767, and FF0964, two sequences were obtained (see Table 1). The PCR product of C3-44 was sequenced directly without prior cloning. Sequence analyses were performed using the programs EuGene (Baylor College of Medicine, Houston, TX), MASE (Faulkner and Jurka, 1988), and DOTS (Kusumi

TABLE 1

HTLV II Specimens Analyzed

Strain	Origin	Material	PCR	PCR strategy	Number of clones sequenced	Source for sequence analysis	Clone designation	GenBank Accession Nos.
C3-44 (MO)	Hairy cell leukemia	Cell culture	Single round	403/404	Direct sequence	Direct sequence	C3-44	U33889
FB330	IVDU/New Orleans	Cell culture	Single round	403/404	1	M13 clone	FB330c1	U33890
LS0567	IVDU/Miami	Cell culture	Single round	403/404	1	M13 clone	LS0567	U33891
DS	Navajo	Uncultured PBMC	Nested	405/406 407/406 408/409	1 (2 halves)[b]	T/A clones	DS	U33892
AP3207	Kayapo	Uncultured WBC[a]	Nested	401/402 403/404	2 (2 base pairs)	T/A clones	AP3207c13	U33893
							AP3207c22	U33894
AP3208	Kraho	Uncultured WBC[a]	Nested	401/402 403/404	2 (2 base pairs)	T/A clones	AP3208c4	U33895
							AP3208c5	U33896
AP3210	Kraho	Uncultured WBC[a]	Nested	401/402 403/404	2 (1 base pair)	T/A clones	AP3210c1	U33897
							AP3210c10	U33898
FF2085	Kaxuyana	Uncultured WBC[a]	Nested	401/402 403/404	2 (4 base pairs)	T/A clones	FF2085c1	U33899
							FF2085c2	U33900
FO567	Guaymi	Uncultured PBMC	Nested	401/402 403/404	1 (1 base pair)	T/A clones	FO567c1	U33901
							FO567c5	U33902
FC561	IVDU/Newark	Cell culture	Single round	403/404	1	M13 clone	FC561c6	U33903
VH	Navajo	Uncultured PBMC	Nested	405/406 407/406 408/409	1 (2 halves)[b]	T/A clones	VH	U33904
LG	Pueblo	Uncultured PBMC	Nested	405/406 407/406 408/409	1 (2 halves)[b]	T/A clones	LG	U33905
SC	Pueblo	Uncultured PBMC	Nested	405/406 407/406 408/409	1 (2 halves)[b]	T/A clones	SC	U33906
AG	Pueblo	Uncultured PBMC	Nested	405/406 407/406 408/409	1 (2 halves)[b]	T/A clones	AG	U33907
CG[c]	Pueblo	Uncultured PBMC	Nested	405/406 407/406 408/409	1 (2 halves)[b]	T/A clones	CG[c]	U33908
BPi	Navajo	Uncultured PBMC	Nested	403/404	1	T/A clone	BPi	U33909
FF0964	Seminole	Uncultured PBMC	Nested	401/402 403/404	2 (identical)	T/A clones	FF0964c1	U33910
							FF0964c2	n/a
FF0754	Seminole	Uncultured PBMC	Nested	401/402 403/404	2 (1 base pair)	T/A clones	FF0754c1	U33911
							FF0754c2	U33912
FF0606	Seminole	Uncultured PBMC	Nested	401/402 403/404	1	T/A clone	FF0606	U33913

[a] WBC, white blood cells; DNA was extracted from a mixture of erythrocytes and leukocytes as described (Maloney *et al.,* 1992).

[b] Sequences were compiled by fusing two half sequences (there were no sequence changes in the short overlap.).

[c] The nucleotide sequence of the *tax/rex* region of sample CG has been reported previously (Hjelle and Chaney, 1992); there is only a 19 base pair overlap with the sequence reported here. n/a, not available.

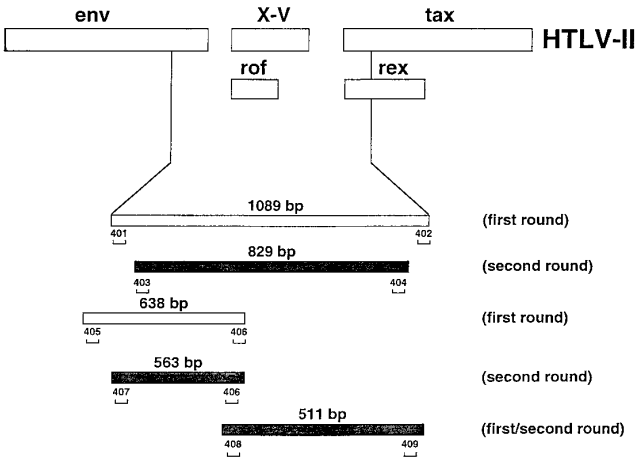

FIG. 1. Genomic location of HTLV-II sequences amplified from infected primary and cultured patient material. Fragment sizes include primer sequences. Shaded fragments were cloned and sequenced.

et al., 1992). Sequences have been submitted to GenBank and accession numbers are listed in Table 1.

Phylogenetic analysis

DNA sequences were aligned using CLUSTAL W (Thompson *et al.,* 1994). The complete alignment was unambiguous, requiring the insertion of only a small number of gaps. Pairwise distances among sequences were estimated using the two-parameter method (Kimura, 1980) to correct for superimposed substitutions. The phylogeny was estimated by applying the neighbor-joining method (Saitou and Nei, 1987) to the pairwise distance matrix. The reliability of clusters within the phylogeny was estimated using 1000 bootstrap analyses. These methods were implemented using CLUSTAL W. Phylogenetic relationships were also analyzed using the maximum parsimony method, implemented using the program DNAPARS from the PHYLIP package (Felsenstein, 1992).

RESULTS

HTLV-II viral sequences were amplified from uncultured peripheral blood of 16 seropositive Amerindians originating from North, Central, and South America as well as from three cultured isolates derived from U.S. intravenous drug users. Using primer pairs and conditions summarized in Table 2, all samples yielded amplification products, although in six cases the region had to be amplified as two separate fragments. In these instances, the two sequences were fused since there were no sequence changes in the short overlap (data not shown).

Figure 2 depicts an alignment of the newly characterized HTLV-II sequences as well as two published sequences of HTLV-II/G12 (Pardi *et al.,* 1993) and HTLV-II/NRA (Lee *et al.,* 1993). HTLV-II/MO is represented by the C3-44 sequence. Nucleotide sequence divergence was assessed by pairwise sequence comparisons and ranged between 0.1 and 6.1%. Translation of each sequence indicated that the majority of HTLV-II strains contained intact (partial) *env,* X-V/*rof,* and (partial) *tax*/*rex* open reading frames. Moreover, splice acceptor sites for the second exons of X-V/*rof* and *tax*/*rex* were highly conserved in all sequences. Only two clones contained in-frame stop codons, one at the carboxy terminus of the *env* gene (AP3207c13) and the other within the *rof* gene (LS0567). Finally, the only length differences were two single base pair deletions in noncoding regions. Taken together, this degree of sequence conservation among geographically diverse viruses further suggests an important role of the X-V/*rof* region (and its gene products) in the HTLV-II life cycle (Ciminale *et al.,* 1992).

Phylogenetic analysis of the newly identified HTLV-II sequences together with previously determined proviral sequences confirmed the presence of two major subtypes of HTLV-II, termed IIa and IIb (Fig. 3). These subtypes were separated by a central branch which was considerably longer than the branches within each subtype. Unless there have been dramatic changes in the rate of evolution in the different lineages, we can assume that the root of the HTLV-II tree lies on this central branch. To determine its position, we selected the nearest available outgroup, which is HTLV-I. Unfortunately, the genomic region examined in this study is not well conserved between HTLV-I and HTLV-II. Therefore, we constructed a phylogenetic tree using *gag* and *pol* protein sequences which are each available for a representative of HTLV-II subtype IIa (MO; Shimotohno *et al.,* 1985) and IIb (NRA; Lee *et al.,* 1993) and which can be easily aligned with the homologous sequences from HTLV-I.

In both trees, we found that the branch from the common ancestor of the HTLV-II sequences to subtype IIa was longer than that to subtype IIb. The tree in Fig. 3 has thus been rooted at a point reflecting the average of the results from the *gag* and *pol* analyses. The longer branch from the root to subtype IIa suggests that, since their divergence, subtype IIa viruses have been evolving faster than viruses belonging to subtype IIb, although this conclusion must remain tentative until a closer outgroup can be used.

As noted previously, almost all HTLV-II strains derived from North and Central Amerindians clustered within HTLV-IIb (lower cluster in Fig. 3). Guaymi and Seminole HTLV-IIb formed individual subclusters supported by high bootstrap values. By contrast, Pueblo and Navajo viruses clustered together and with a virus isolated from an American drug user (FC561). Also positioned in IIb was HTLV-II/NRA, a virus identified in a Caucasian male with an atypical hairy cell leukemia (Lee *et al.,* 1993).

Viruses infecting the Brazilian Amerindians clustered within HTLV-IIa (upper cluster), together with two IVDU

TABLE 2

PCR Primer Pairs Used for Amplification of HTLV-II Subgenomic Fragments

Primer pairs	Designation	Nucleotide sequence	Location[a]	Fragment size[b]	Genomic region
Outer pair	401	5′-CTCCTATTCTGGGAACAAGGGGGTTT-3′	6317–6342	1089 bp	*env/rof/tax*
	402	5′-GAGCCGATAACGCGTCCATCG-3′	7386–7406		
Inner pair	403	5′-GGCTGGGGACTAAACTGGGATCCTGG-3′	6443–6468	829 bp	*env/rof/tax*
	404	5′-CCAAACACGTAGACGGGGGATCC-3′	7249–7271		
Outer pair	405	5′-GAGGTTGACAAAGACATCTCCC-3′	6212–6233	638 pb	*env/rof*
	406	5′-CTTCAGGGTTATGTGGATTT-3′	6828–6847		
Inner pair	407	5′-TATGCAGCCCAAAATAGACGAGG-3′	6287–6309	563 bp	*env/rof*
	406	5′-CTTCAGGGTTATGTGGATTT-3′	6828–6847		
Outer pari	408	5′-TAACCCCCCGCTCACATTCCTCC-3′	6781–6302	511 bp	*rof/tax*
	409	5′-CAATCGGCCTGTACACAATC-3′	7273–7292		
Inner pair	408	5′-TAACCCCCCGCTCACATTCCTCC-3′	6781–6302	511 bp	*rof/tax*
	409	5′-CAATCGGCCTGTACACAATC-3′	7273–7292		

[a] Primer pairs were numbered according to the published sequence of HTLV-II/MO (Shimotohno *et al.,* 1985).
[b] Reaction conditions were 1 min, 92°; 1.5 min, 50°; 1 min, 60°; 30 cycles.

strains, a strain from a patient with hairy cell leukemia, and a single Navajo strain. The Brazilian viruses, however, formed a distinct subcluster (supported by high bootstrap values) and were separated from the other IIa viruses by a relatively long branch. Within that subcluster, viruses from three different Brazilian tribes (Kraho, Kayapo, and Kaxuyana) were very closely related, and the two viruses from one tribe (Kraho) were not significantly more related to each other than to those from the two other tribes (compare Fig. 2). This is of interest given the considerable geographic distance between these tribes.

DISCUSSION

In this paper we present the first comprehensive phylogenetic analyses of HTLV-II sequences derived from several endemically infected Amerindian populations. Comparing viral sequences from 16 different individuals representing North, Central, and South American Indians, including members of three different remote Amazonian tribes, we found that Amerindian HTLV-II strains clustered within the same two sequence subtypes (designated IIa and IIb) previously identified for North American drug users (Hall *et al.,* 1993; Switzer *et al.,* 1995a; Seleni *et al.,* 1995).

Among the Amerindians outside of Brazil, the subtype was largely IIb, as has been reported in other studies (Dube *et al.,* 1993; Hjelle *et al.,* 1993; Switzer *et al.,* 1995b). Other than in the Brazil Amerindians, the only IIa strains were found among Amerindians living in the southwestern United States, among whom drug abuse exposure could not be excluded (Hjelle *et al.,* 1993). However, as we first presented in 1993 (Biggar *et al.,* 1993), the Amerindians of the Amazon basin in Brazil had subtype IIa. Furthermore, HTLV-IIa sequences from these Indians formed their own distinct subcluster within subtype IIa, indicating that they were not the source of HTLV-IIa in North American drug users.

The clustering of Amerindian-derived HTLV-II sequences within two distinct subtypes (IIa and IIb) was also described in recent studies by Switzer *et al.* (1995a) who used LTR sequences for subtype determination, and Ishak *et al.* (1995) who used *env* sequences. In addition, these investigators reported that HTLV-II sequences from Kayapo Indians clustered distinct from HTLV-II strains found in the U.S. and European drug users. Although our results are in general agreement with these two studies, we interpret certain aspects of the HTLV-II phylogeny differently. For example, the position of the root in our phylogeny (Fig. 3) strongly contradicts the conclusions of Switzer *et al.* (1995a), who placed the root deep within subtype IIb. In their phylogenetic tree, the root separated a single subtype IIb virus (strain JG) from all other subtype IIb and subtype IIa sequences, such that these authors conclude that subtype IIa evolved from subtype IIb. This scenario seems most unlikely. In their study, Switzer *et al.* (1995a) specifically chose to examine LTR sequences because they are more variable than other regions of the HTLV-II genome. They also used HTLV-I as an outgroup to root their tree. However, the position of this root is likely to be unreliable because the LTR sequences of HTLV-I and HTLV-II are too divergent to be suitable for this analysis. Furthermore, the JG sequence (from an individual in New York) appears to be identical to (i.e., has a zero branch length from) the inferred sequence at the root of their tree. Since it is believed that the ancestral virus probably existed tens of thousands of years ago, it seems quite unlikely

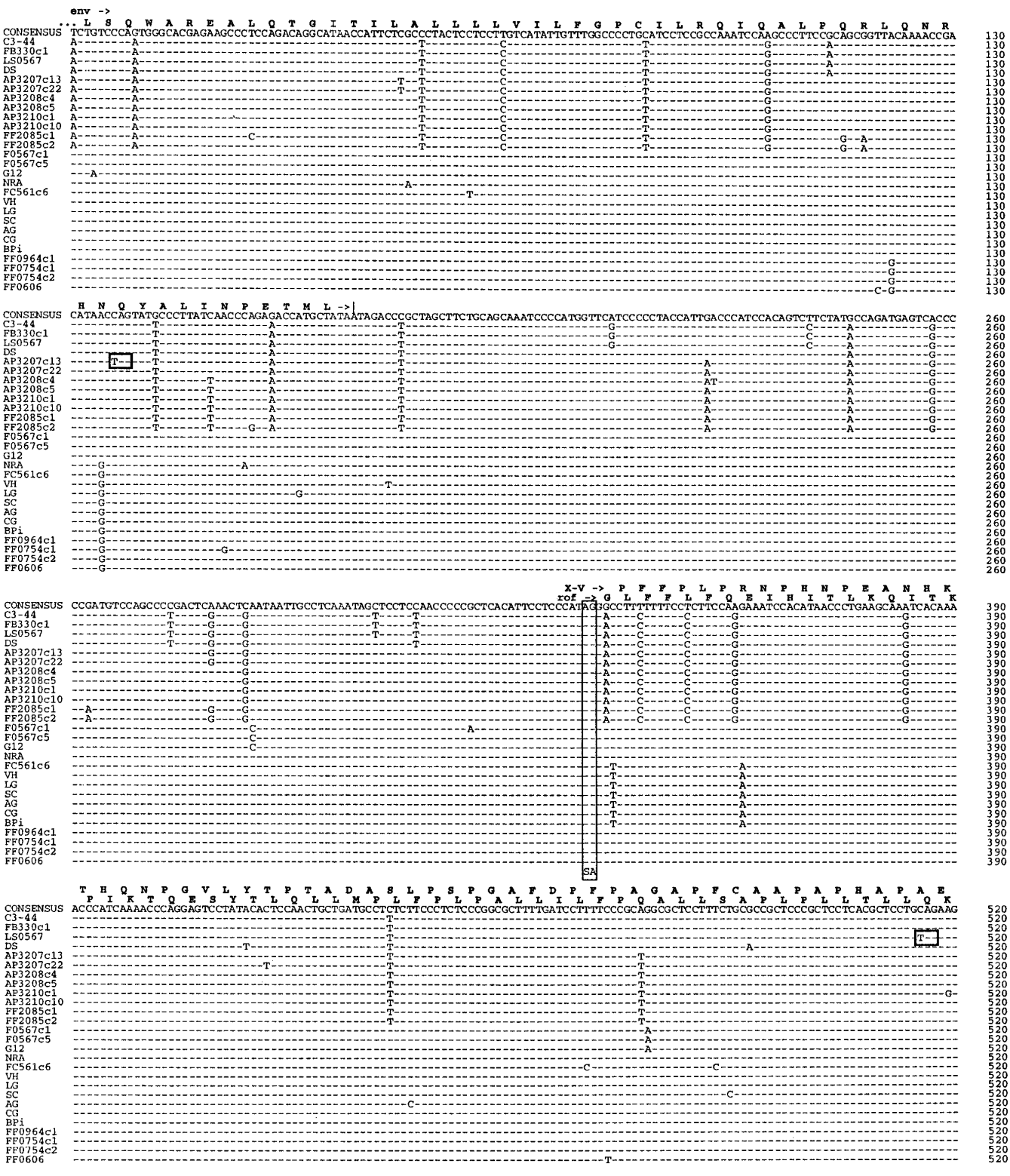

FIG. 2. Nucleotide sequence alignment of PCR-derived HTLV-II sequences. Nucleotide sequences are compared to a consensus sequence generated by MASE (Faulkner and Jurka, 1988). Dashes denote sequence identity with the consensus sequence and dots represent gaps in the alignment. Deduced amino acid sequences of *env,* X-V/*rof, tax*/*rex* (translated from the consensus sequence) are indicated. The splice acceptor sites for X-V/*rof* and *tax*/*rex* are boxed. Also boxed are the in-frame stop codons in AP3207c13 and LS0567. Two single base pair deletions are circled.

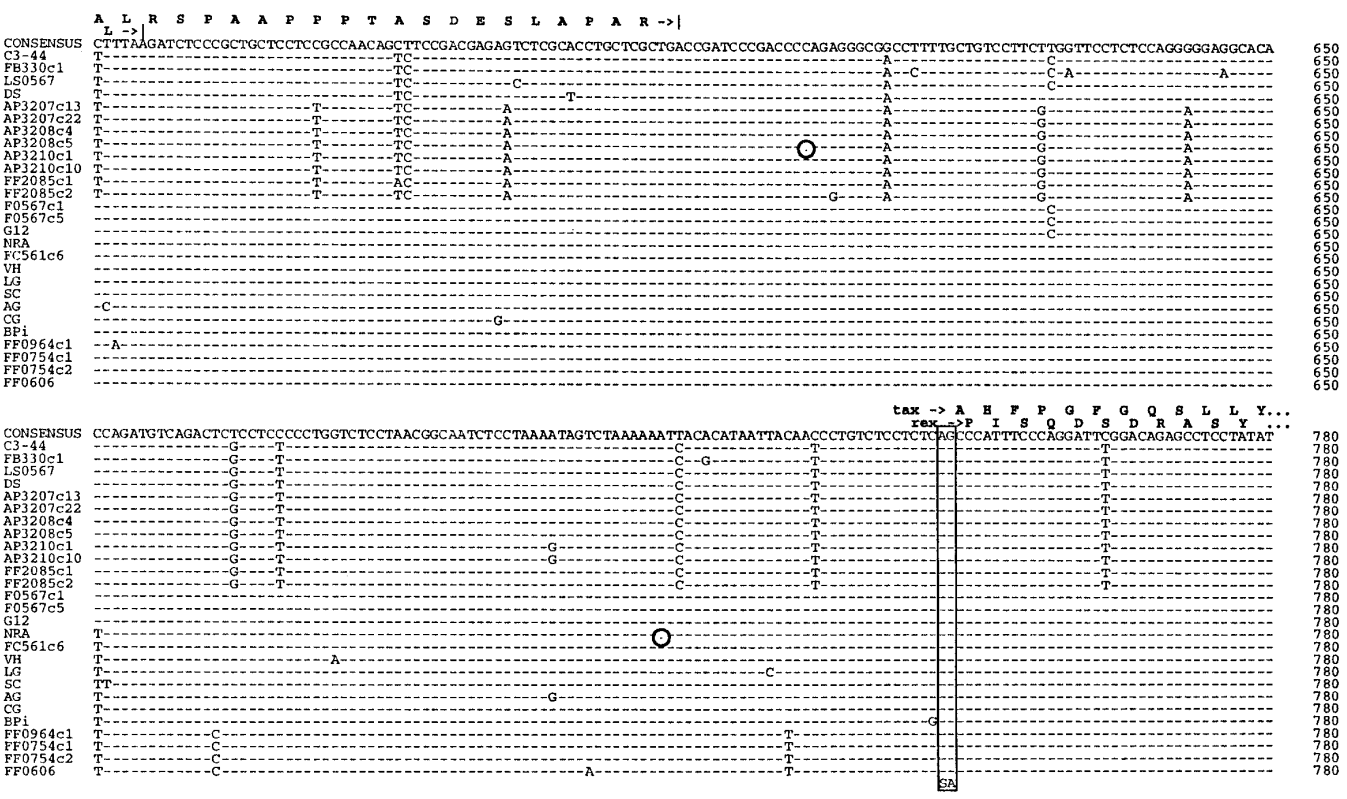

FIG. 2—*Continued*

that an individual from New York could have acquired such a "fossil" virus.

Finally, Ishak *et al.* (1995) reported that the HTLV-II strains infecting Kayapo Indians belong to a distinct phylogenetic group and proposed to call this lineage HTLV-IIc. Although we agree that HTLV-II strains infecting Brazilian Amerindians form a distinct phylogenetic lineage, we feel that the designation HTLV-IIc is inappropriate because it implies equidistance to HTLV-IIa and IIb. Instead, our analysis of viruses from three different Amazonian Indian tribes indicates that these viruses form a subcluster within subtype IIa.

The phylogenetic relationships of the different HTLV-II strains in Fig. 3 also raise questions concerning the origin of HTLV-II and its introduction into New World Indian populations. As we have pointed out (Neel *et al.,* 1994), the widespread distribution of endemically infected Amerindians suggests that the original migrants to the New World more than 15,000 years ago brought HTLV-II with them. However, the occurrence of both IIa and IIb suggests that the migrating populations brought at least two subtypes of HTLV-II into the New World. We were unable to find HTLV-II (or HTLV-I) in indigenous populations of eastern Siberia (Neel *et al.,* 1994), and Batsuuri *et al.* (1993) were similarly unable to find evidence of HTLV-I in Mongolia. However, Hall *et al.* (1994b) recently reported finding HTLV-II in Mongolia, allowing us to speculate that migration to the New World originated from this population (a pattern also supported by genetic data (Neel *et al.,* 1994)). According to Hall *et al.* (1994b), the subtype in Mongolia is IIa, leaving the origin of Amerindian IIb still open to question. In addition to its presence in IVDUs, endemic IIb has also been reported to infect African Pygmies (Gessain *et al.,* 1995), but, once again, Pygmies are an unlikely source for infecting either IVDUs or Amerindian indigenous groups. A virus related to HTLV-II has recently been identified in bonobos (pygmy chimpanzees) (Vandamme *et al.,* 1995), and there is evidence for cross-species transmission of HTLV-I from nonhuman primates to humans (Koralnik *et al.,* 1994). Thus, there may be additional HTLV-II types in humans that have not yet been identified.

The worldwide distribution of the two HTLV-II subtypes in remote and unrelated populations suggests genetic stability of HTLV-II for many millennia, possibly ever since man emerged from Africa some 100,000 years ago. Despite this antiquity, there is relatively little strain variation even in different populations with no contact over very long periods of time. This lack of variation is in marked contrast to the variations seen in HIV-1 and HIV-2 (Sharp *et al.,* 1994), pathogens only recently epidemic in humans. Adverse health effects associated with HTLV-II infection have not yet been well documented, but if they exist and are found to be restricted to some subpopula-

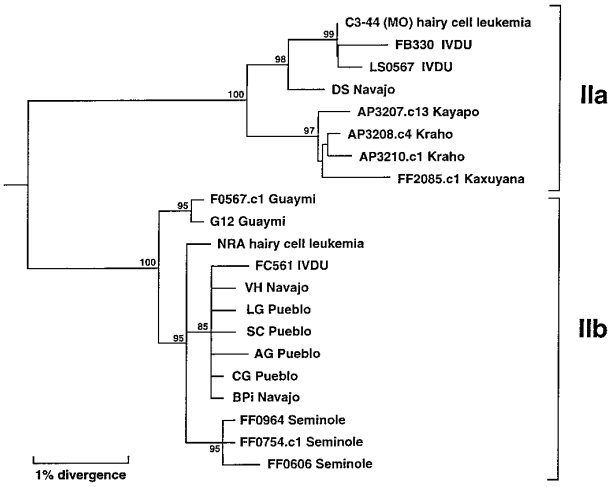

FIG. 3. Phylogenetic relationships among HTLV-II strains. Subtype IIa and IIb represent the upper and lower clusters, respectively. The tree was derived from 779 aligned nucleotide sites in the X-V/*rof* region of the HTLV-II genome. Only one sequence per HTLV-II strain was used for analysis. Horizontal branch lengths are drawn to scale (the bars represent 0.01 nucleotide substitutions per site, or 1% divergence); vertical separations are for clarity only. Numbers at nodes indicate the percentage of 1000 boot-straps in which the cluster to the right is supported (only values >80% are shown). The position of the root is approximate (see text). Phylogenetic relationships were also analyzed by the maximum parsimony method (Felsenstein, 1992), which identified 100 equally parsimonious tress. A consensus tree derived from these contained all of the significant clusters seen in the neighbor-joining tree values (i.e., those supported by high bootstrap) and differed only in placing NRA closer to the Navajo/Pueblo cluster than to the Seminole cluster.

tions, then the differentiation between strains may help to clarify regions of HTLV-II that are important for pathogenicity. Alternatively, because of a very ancient virus/host relationship, HTLV-II may have adapted to persist in humans without causing serious health problems.

Finally, high prevalence HTLV-II seems to be endemically transmitted only among small genetically homogeneous tribes but not in large heterogeneous populations. One explanation might be that HTLV-II is preserved in these small groups only because of heavy dependence on breast feeding, including a pattern of communal breast feeding in which any lactating woman may breast feed any infant in the village (Black *et al.,* 1994). However, for most of mankind's history, breast feeding has been the only source of infant nutrition and that remains true for many outbred populations, which should then also be HTLV-II infected. Yet the prevalence of HTLV-II in outbred populations of Africa, Asia, and Europe is very low. An alternative suggestion is that transmission may be facilitated by genetic factors. For example, HLA similarity will be more homogeneous in small in-bred groups but will vary greatly between individuals of outbred populations. One could speculate that HLA type might be less important for transmission than the fact of homogeneity between the infected and exposed persons.

ACKNOWLEDGMENTS

We thank J. B. Wilson for the artwork. These studies were supported by the Sequence Analysis Core of the Birmingham Center for AIDS Research (P30-AI-27767) and the Birmingham Veterans Administration Medical Center, Birmingham, Alabama.

REFERENCES

Batsuuri, J., Dashnyam, B., Maidar, J., Battulga, D., Dorjsuren, D., and Ishida, T. (1993). Absence of human T-lymphotropic retrovirus type-1 (HTLV-1) in different populations of Mongolia. *Scand. J. Infect. Dis.* 25, 398–399.

Biggar, R. J., Hahn, B., Taylor, M., Black, F., Sukarnik, R., and Neel, J. (1993). Viral Archeology of HTLV. *IX Int. AIDS Conference.* Abstract PO-CO6-2715. Berlin, June 6–11, 1993.

Black, F. L., Biggar, R. J., Neel, J. V., Maloney, E. M., and Waters, D. J. (1994). Endemic transmission of HTLV type II among Kayapo Indians of Brazil. *AIDS Res. Hum. Retroviruses* 10, 1165–1171.

Ciminale, V., Pavlakis, G. N., Derse, D., Cunningham, C. P., and Felber, B. K. (1992). Complex splicing in the human T-cell leukemia virus (HTLV) family of retroviruses: Novel mRNAs and proteins produced by HTLV type I. *J. Virol.* 66, 1737–1745.

Dube, D. K., Sherman, M. P., Saksena, N. K., Bryo-Gornia, V., Mendelson, J., Love, J., Arnold, C. B., Spicer, T., Dube, S., Glaser, J. B., Williams, A. E., Nishimura, M., Jacobson, S., Ferrer, J. F., Del Pino, N., Quiruelas, S., and Poiesz, B. J. (1993). Genetic heterogeneity in human T-cell leukemia/lymphoma virus type II. *J. Virol.* 67, 1175–1184.

Faulkner, D. M., and Jurka, J. (1988). Multiple aligned sequence editor (MASE). *Trends Biochem. Sci.* 13, 321–322.

Felsenstein, J. (1992). PHYLIP (Phylogeny Inference Package), 3.5c ed. Department of Genetics, University of Washington, Seattle, Washington.

Gessain, A., Mauclere, P., Froment, A., Biglione, M., Lettesran, J. Y., Tekaia, F., Millan, J., and de The, G. (1995). Isolation and molecular characterization of a human T lymphotropic virus type II, subtype B, from a healthy Pygmy in a remote area of Cameroon: An ancient origin for HTLV-II in Africa. *Proc. Natl. Acad. Sci. USA* 92, 4041–4045.

Goubau, P., Desmyter, J., Ghesquiere, J., and Kasereka, B. (1992). HTLV-II among Pygmies. *Nature* 359, 201.

Hall, W. W., Takahashi, H., Liu, C., Kaplan, M. H., Scheewind, O., Ijichi, S., Nagashima, K., and Gallo, R. C. (1993). Multiple isolates and characteristics of T-cell leukemia virus type II. *J. Virol.* 66, 2456–2463.

Hall, W. W., Kubo, T., Ijichi, S., Takahashi, H., and Zhu, S. H. (1994a). Human T cell leukemia/lymphoma virus type II (HTLV-II): Emergence of an important newly recognized pathogen. *Semin. Virol.* 5, 165–178.

Hall, W. W., Zhu, S. W., Horal, P., Furuta, Y., Zagaany, G., and Vahina, A. (1994b). HTLV-II infection in Mongolia. *AIDS Res. Hum. Retroviruses* 10, 443.

Hjelle, B., and Chaney, R. (1992). Sequence variation of functional HTLV-II *tax* alleles among isolates from an endemic population: Lack of evidence of oncogenic determinants in *tax. J. Med. Virol.* 36, 136–141.

Hjelle, B., Zhu, S. W., Takahashi, H., Ijichi, S., and Hall, W. W. (1993). Endemic human T cell leukemia virus type II infection in southwestern U.S. Indians involves two prototypic variants of virus. *J. Infect. Dis.* 168, 737–740.

Ishak, R., Harrington, W. J., Azevedo, V. A., Eiraku, N., Ishak, M. O. G., Guerreiro, J. F., Santos, S. B., Kubo, T., Monken, C., Alexander, S., and Hall, W. W. (1995). Identification of human T cell lymphotropic virus type IIa infection in the Kayapo, an indigenous population in Brazil. *AIDS Res. Hum. Retroviruses* 11, 813–821.

Kimura, M. (1980). A simple method for estimating evolutionary rates

of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111–120.

Koralnik, I. J., Boeri, E., Saxinger, W. C., Monico, A. L., Fullen, J., Gessain, A., Guo, H-G., Gallo, R. C., Markham, P., Kalyanaraman, V., Hirsch, V., Allan, J., Murthy, K., Alford, P., Slattery, J. P., O'Brien, S. J., and Franchini, G. (1994). Phylogenetic associations of human and simian T-cell leukemia/lymphotropic virus type I strains: Evidence for interspecies transmission. *J. Virol.* 68, 2693–2707.

Kusumi, K., Conway, B., Cunningham, S., Berson, A., Evans, C., Iversen, A. K. N., Colvin, D., Gallo, M. V., Coutre, S., Shpaer, E. G., Faulkner, D. V., deRonde, A., Volkman, S., Williams, C., Hirsch, M. S., and Mullins, J. I. (1992). Human immunodeficiency virus type 1 envelope gene structure and diversity in vivo and after cocultivation in vitro. *J. Virol.* 68, 875–885.

Lairmore, M. D., Jacobson, S., Garcia, F., De, B. K., Castillo, L., Larreategui, M., Roberts, B. D., Levine, P. H., Blattner, W. A., and Kaplan, J. E. (1990). Isolation of human T lymphotropic virus type 2 from Guaymi Indians in Panama. *Proc. Natl. Acad. Sci. USA* 87, 8840–8844.

Lee, H., Idler, K. B., Swanson, P., Aparicio, J. J., Chin, K. K., Lax, J. P., Nguyen, M., Mann, T., Leckie, G., Zanetti, A., Marinucci, G., Chen, I. S. Y., and Rosenblatt, J. D. (1993). Complete nucleotide sequence of HTLV-II isolate NRA: Comparison of envelope sequence variation of HTLV-II isolates from U.S. blood donors and U.S. and Italian IV drug users. *Virology* 196, 57–69.

Levine, P. H., Jacobson, S., Elliott, R., Cavallaro, A., Colclough, G., Dorry, C., Stephenson, C., Knigge, R. M., Nishimura, M., Taylor, M., Wiktor, S., and Shaw, G. M. (1993). HTLV-II infection in Florida Indians. *AIDS Res. Hum. Retroviruses* 9, 123–127.

Maloney, E. M., Biggar, R. J., Neel, J. V., Taylor, M. E., Hahn, B. H., Shaw, G. M., and Blattner, W. A. (1992). Endemic human T-cell lymphotropic virus type II infection among isolated Brazilian Amerindians. *J. Infect. Dis.* 166, 100–107.

Neel, J. V., Biggar, R. J., and Sukernik, R. (1994). Virologic and genetic studies relate Amerind origins to the indigenous people of Mongolia/Manchuria/southeastern Siberia region. *Proc. Natl. Acad. Sci. USA* 91, 10737–10741.

Pardi, D., Switzer, W. M., Hadlock, K. G., Kaplan, J. E., Lal, R. B., and Folks, T. M. (1993). Complete nucleotide sequence of an Amerindian human T-cell lymphotropic virus type II (HTLV-II) isolate: Identification of a variant HTLV-II subtype B from a Guaymi Indian. *J. Virol.* 67, 4659–4664.

Poiesz, B. J., Sherman, M. P., Saksena, N. K., Dube, D., Dube, S., Gavalchin, J., Fan, N., Lane, M., and Paul, B. (1993). The biology and epidemiology of the human T cell lymphoma/leukaemia viruses. *In* "Frontiers of Infectious Diseases: Focus on HIV" (H. C. Neu, J. C. Levy, and R. A. Weiss, Eds.), pp 189–205. Churchill Livingston, London.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Seleni, M., Cattaneo, E., Casoli, C., and Bertazzoni, U. (1995). Identification of IIa and IIb molecular subtypes of human T cell lymphotropic virus type II among Italian injecting drug users. *J. Acq. Immuno. Defic. Retrovirol.* 15, 516–520.

Sharp, P. M., Robertson, D. L., Gao, F., and Hahn, B. H. (1994). Origins and diversity of human immunodeficiency viruses. *AIDS* 8, S27–S42.

Shimotohno, K., Takahashi, Y., Shimizu, N., Gojobori, T., Golde, D. W., Chen, I. S. Y., Miwa, M., and Sugimura, T. (1985). Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type II: An open reading frame for the protease gene. *Proc. Natl. Acad. Sci. USA* 82, 3101–3105.

Switzer, W. M., Pieniazek, D., Swanson, P., Samdal, H. H., Soriano, V., Khabbaz, R. F., Kaplan, J. E., Lal, R. B., and Heneine, W. (1995a). Phylogenetic relationship and geographic distribution of multiple human T-cell lymphotropic virus type II subtypes. *J. Virol.* 69, 621–632.

Switzer, W. M., Black, F. L., Pieniazek, D., Biggar, R. J., Lal, R. B., and Heneine, W. (1995b). Endemicity of a unique human T cell lymphotropic virus type II-subtype in the Kayapo Indians of Brazil. *AIDS Res. Hum. Retroviruses,* in press.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W — Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.

Vandamme, A-M., Liu, H-F., Van Brussel, M., De Meurichy, W., Desmyter, J., and Goubau, P. (1995). The presence of a divergent T-lymphotropic virus in a wild-caught *Pan Paniscus* supports an African origin for the HTLV/STLV group of viruses. *J. Gen. Virol.,* in press.